A new DEM extraction method for Hexagon spy imagery and application to Bhutan glaciers

Josh Maurer, Summer Rupper

Department of Geological Sciences, Brigham Young University

ABSTRACT - Declassified Hexagon stereo spy images have near-global coverage extending back to the 1970's, yet remain a largely untapped resource for land change studies. Unavailable satellite exterior orientation data for these images make digital elevation model (DEM) extraction difficult in terms of time and accuracy. A new automated workflow for DEM extraction is presented that eliminates the need for manual ground control point selection. The method is applied to reconstruct a DEM from 1974 imagery over a large glacierized region in the Bhutan Himalayas. Glacier changes over several decades are visualized using a DEM differencing method. These results demonstrate the value of Hexagon imagery when applied to land change studies.

INTRODUCTION

A digital elevation model (DEM) is a 3D representation of a terrain surface. DEMs are used in countless applications, such as hydrological/mass movement modeling or even 3D visualizations in flight simulators (Betts and DeRose, 1999; Huggel et al., 2008; Perry, 2004). One commonly employed technique often used for land-change studies is known as DEM differencing (James et al., 2012; Kucera, 1992), which compares DEMs over the same region from different time periods. This allows quantification of surface elevation changes due to erosion, landslides, earthquakes, melting glaciers, construction of man-made features, and many other factors. It follows that historical DEMs are useful for land-surface change studies.

The Hexagon image database contains thousands of satellite images covering large regions of the globe (one image covers approximately 3400 km^2) at 6 to 9 meter ground resolution, acquired between 1971 and 1986. These images could potentially be of immense value for land-change studies, as overlapping camera images allow for stereo matching and DEM extraction. A new workflow is presented to efficiently and accurately extract DEMs and orthorectified imagery without the need for manual ground control point selection, rendering the Hexagon database much more accessible over a wide variety of disciplines.

BACKGROUND

Photogrammetry

Various methods exist for obtaining DEMs, such as Lidar, Synthetic Aperture Radar, land surveying, and photogrammetry. It is the photogrammetric method (applied to the historical Hexagon stereo imagery) that is the focus of this study. The basic principle behind the technique is quite simple. Two

or more images are taken of a terrain surface from separate locations, and light rays projected from the camera optical centers will intersect at a point in space (figure). This is analogous to human vision with the left and right eyes providing a perceived sense of depth.

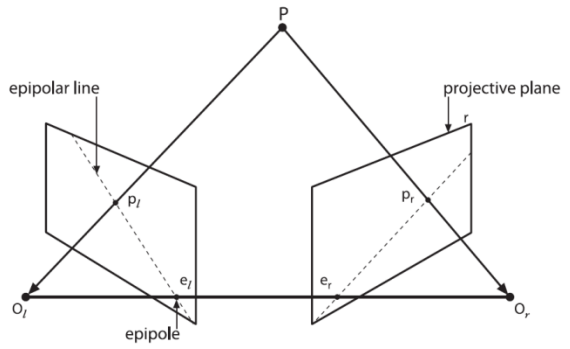

Figure 1. Stereo geometry. O_l and O_r are the centers of projection for the left and right images, respectively. P is the observed point, projected onto the image planes at p_l and p_r (Hartley and Zisserman, 2003).

KH-9 Hexagon mapping program

Archival satellite images provide an important worldwide record of land-surface change. The Keyhole-9 (codename Hexagon) program consisted of photographic reconnaissance satellite systems operational from 1971 to 1986. Thousands of photographs worldwide were acquired by the "mapping camera" system at a resolution of 9 meters (improved to 6 in later missions) with near global coverage (Surazakov, 2009). Since the images could be of historical value for global change research and were no longer critical to national security, the collection was made available to the public in 2002. The U.S. Geological Survey (USGS) then used high performance photogrammetric film scanners to create digital products at 7-14 micron resolution; some of these images are available for free download, others require a nominal fee.

Challenges with Hexagon imagery

Most of the mission-related documentation of the Hexagon program remains classified (Surazakov, 2009). This includes any ephemeris data such as satellite exterior orientation parameters. Thus it is necessary to manually obtain/select ground control points, through GPS or viewing modern high-resolution georeferenced imagery. Corresponding points must then be identified in the historical image of interest. After these points are identified, the orientation of the satellite can be estimated via the collinearity condition (Wolf and Dewitt, 2000). However, the manual selection of ground control points for historical images is a tedious and time-consuming work. The process is made difficult by the passing of several decades between image acquisition dates (i.e. between the historical and modern reference images). Commonly used control point markers such as road intersections, corners of buildings, stream intersections, etc. have often undergone significant change. In the case of remote study areas, man-made structures are rare. Furthermore, high erosion rates, temporal variability in snow cover, cloud cover, etc. make accurate identification of natural features very difficult. Ironically, these remote regions of the world could benefit most from utilizing historical imagery for landscape evolution, as direct observation via field studies has not occurred. A prime example is that of the Bhutanese Himalayan region, where multi-decadal changes in the extent of glacial ice

and glacial lakes are of interest to better understand the effect of dwindling glacial ice on water resources. In regions such as this, Hexagon images could be of immense value.

This work will present a solution to the problem of estimating camera orientation when overlapping images (i.e. stereo vision) are available, yet accurate ground control point selection is not feasible. A digital elevation model is extracted over a key region in the Bhutan Himalayas, highlighting the usefulness of these historical spy satellite images for land-change studies.

METHODS

The workflow is close to being fully automated, with a few initial user inputs needed to get the process started. It is implemented within the MATLAB programming environment.

Image Preprocessing

A few preliminary processing steps must be taken before the images are suitable for stereo matching. First regularly spaced reseau marks are detected using a moving window that computes the local standard deviation around each pixel. Assuming regular grid spacing, the reseau locations are used to correct any geometric image distortions that may have occurred during 4 decades of storage, film scanning process, etc. Also, a locally adaptive filter is applied to enhance image contrast (Surazakov, 2009).

Stereo Rectification and Matching

An ideally calibrated stereo imaging system has both image planes perfectly aligned. However, in practice this is rarely the case. Thus, before computing a stereo disparity map, Hexagon image pairs are rectified so features in both images appear on the same horizontal rows. The rectification can be thought of as rotating the image planes around their optical centers until focal planes become coplanar (Fusiello and Irsara, 2008). Features in both images are detected using SURF descriptors (Bay et al., 2006) then matched using normalized cross correlation (Lewis, 1995). The matched points are subsequently used to compute the epipolar geometry relating the two images, known as the fundamental matrix. Any outlying matches inconsistent with the epipolar geometry are rejected using the RANSAC technique, and the remaining points are used to compute a homography transformation, effectively aligning features in both images (Figure 2).

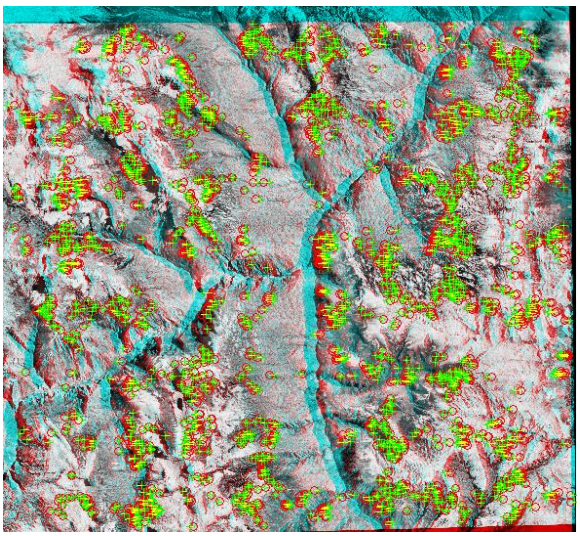

Figure 2. A composite red-cyan anaglyph image displaying the rectified Hexagon stereo pair. The left image is red and the right image is cyan. SURF features matched using normalized cross correlation are shown as red "o" symbols (left image) and green "+" symbols (right image).

An implementation of the semi-global block matching algorithm (Hirschmuller, 2008) is utilized from the open source software package OpenCV to compute the stereo disparity map (Figure 3). Subsequently, matched pixels are projected back to their respective pre-rectified image coordinates using the inverse of the homography transformation.

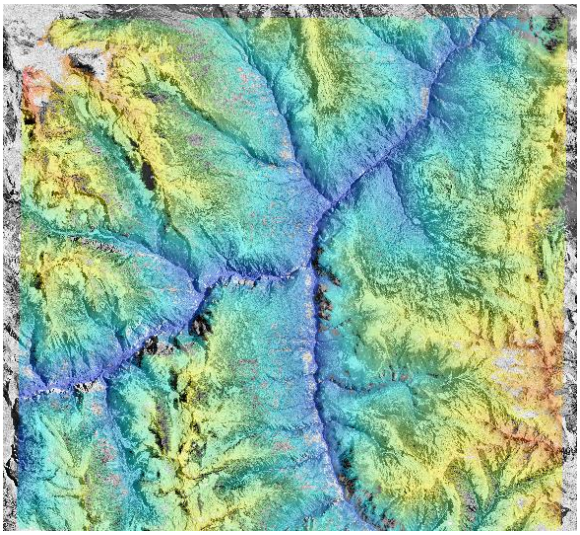

Figure 3. Stereo disparity map computed using the semi-global block matching algorithm. "Cool" pixels represent smaller disparities (valleys further from camera), while "hot" pixels represent larger disparities (mountains closer to camera).

Estimating Satellite Exterior Orientation

Elevation models already exist for most of the earth's surface, the prime example being the freely available Shuttle Radar Topography Mission (SRTM) global DEM from the year 2000. The following section describes a way of using a reference DEM to back out the orientation of the Hexagon satellite at the time of image acquisition. Essentially, the reference DEM pixels are projected onto an image plane twice, before and after a slight shift. This effectively creates an "artificial" stereo disparity map.

This artificial disparity map is then compared to the actual disparity map (computed from the Hexagon images) using a measure of similarity. Parameters describing the relative orientation between the camera and the reference DEM are then optimized using a nonlinear solver method. When the two disparity maps match, the relative orientation between the reference DEM and camera is known, hence the exterior orientation of the satellite imaging system is also known.

The orientation (pose) of an object can be described by three rotation angles, $\omega, \varphi,$ and κ, around the $X, Y,$ and Z axes, respectively. These three angles are combined in a rotation matrix R:

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & -\sin\omega \\ 0 & \sin\omega & \cos\omega \end{bmatrix}. \quad (1)$$

$$R_y = \begin{bmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{bmatrix}. \quad (2)$$

$$R_z = \begin{bmatrix} \cos\kappa & -\sin\kappa & 0 \\ \sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3)$$

$$R = R_x R_y R_z \,. \quad (4)$$

The location of an object is defined by three translations t_x, t_y, and t_z in three directions. These are composed into vector t:

$$t = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}. \quad (5)$$

Projection of a 3D object onto an image plane is described by the camera matrix K,

which is comprised of the focal length in horizontal pixel units (f_u), vertical pixel units (f_v), and principal point (c_u, c_v):

$$K = \begin{bmatrix} f_u & 0 & c_u \\ 0 & f_v & c_v \\ 0 & 0 & 1 \end{bmatrix} . \quad (6)$$

Lastly, object coordinates are defined by vector M, and image coordinates by vector m:

$$M = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} . \quad (7)$$

$$m = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} . \quad (8)$$

Thus, the projection of any point in 3D space onto an image plane can be defined as:

$$s\, m = A\, [R|t]\, M . \quad (9)$$

where s is a scale factor (Hartley and Zisserman, 2003).

By slightly shifting the reference DEM horizontally, an artificial disparity is induced. This simulates viewing the DEM from two locations. Both real and artificial disparity maps are converted to grayscale images (i.e. normalized). Thus, only relative pixel intensity values are used in the optimization. The downhill simplex (Nelder-Mead) method is used to minimize an error function S:

$$S = \sum_{i=1}^{m} (r_i - a_i)^2 . \quad (10)$$

where m is the number of pixels in the disparity map, and the variables r_i and a_i represent grayscale pixel values at single pixel locations (u, v) in the real and artificial disparity maps, respectively.

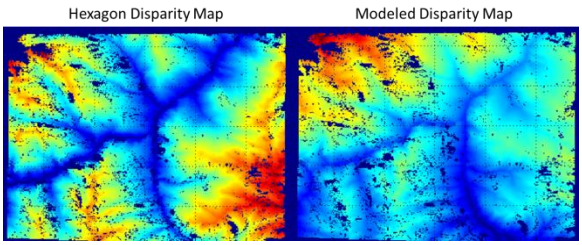

Figure 4. Left-Hexagon disparity map computed using the semi-global block matching algorithm. Right-Modeled disparity map constructed by projecting points from the SRTM reference DEM onto an image plane twice (before and after a slight shift in position), then computing the pixel distance between respective points. This shows the initial guess from the POSIT algorithm. Note the modeled map does not match the actual Hexagon map, as rotation and translation parameters have not yet been optimized. "Cool" pixels represent smaller disparities (valleys further from camera), while "hot" pixels represent larger disparities (mountains closer to camera).

An initial guess for the nonlinear solver is obtained using the POSIT algorithm with image corner coordinates as input (Dementhon and Davis, 1995). Subsequently, in each iteration of the solver routine, 3D points from the reference DEM are translated and rotated according to the six parameters, then projected onto the image plane twice (as noted above, the reference DEM is slightly shifted horizontally to induce an artificial disparity) for both the pre and post-shifted DEM. The quantity a_i is the distance between the two projected points on the image plane in pixel units, normalized to grayscale. Since only the normalized pixel intensity values are used in the optimization, the distance which the reference DEM is shifted is arbitrary. The solver is allowed to proceed until a certain threshold is reached (user-specified minimum tolerance or step size). Also note

that the image plane is defined at the origin with zero rotation (i.e. the matrix $[I|0]$).

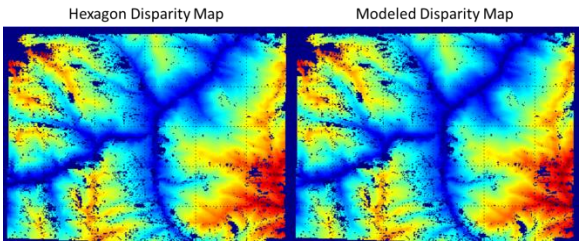

Figure 5. Left-Hexagon disparity map computed using the semi-global block matching algorithm. Right-Modeled disparity map after optimization. Note that the two maps now match. See Figure 4 caption for more details.

Following the pixel intensity-based registration, the solution is further refined through bundle adjustment (Triggs et al., 2000), and the points are triangulated in 3D space using the direct linear method (Hartley and Zisserman, 2003). As a final step, a variant of the iterative-closest-point (ICP) algorithm is used to register the triangulated point cloud to the reference DEM surface. It varies from traditional ICP in that it includes a scale parameter in addition to the rigid rotation and translation. The scale parameter allows for correction of error induced by slightly inaccurate focal point or pixel resolution measurements via nonlinear optimization. Only points on stable terrain are used for this registration. Any points located on known glacial ice, recent landslide debris, etc. are excluded to ensure accuracy. Linear interpolation is then used to construct a DEM surface from the registered 3D points (Figure 6).

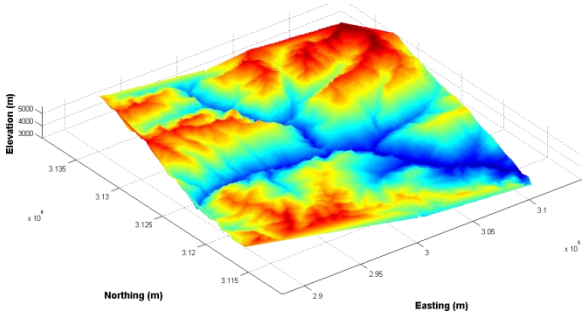

Figure 6. Example of a finished Hexagon DEM from images acquired on January 2nd, 1974.

RESULTS

The mountainous study region in the Bhutan Himalayas (Figure 7) represents an extreme end member with regard to terrain roughness. This in combination with large areas of low image contrast over glaciated and snow-covered regions make accurate DEM extraction particularly difficult.

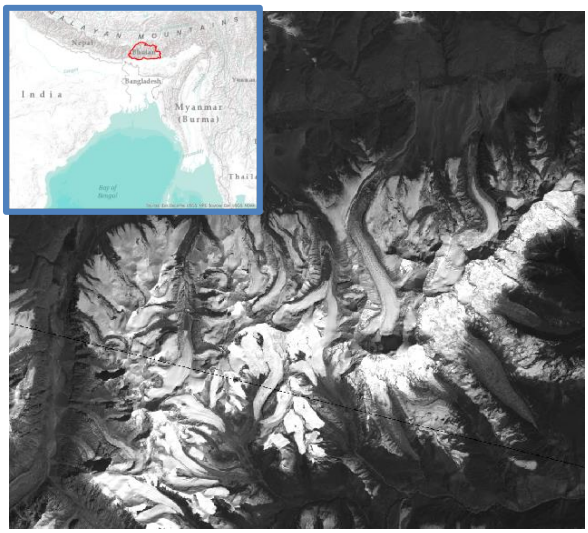

Figure 7. Landsat panchromatic image showing study region in Bhutan/China, with upper left inset showing the Kingdom of Bhutan in red outline. Temporal changes in area and volume of glaciers and lakes are of interest in this region to better understand the effect of dwindling glacial ice on water resources.

Nevertheless, the method described above is applied to a pair of Hexagon images acquired on January 2nd, 1974, and a DEM

is computed for the region at 15 meter resolution. It is then subtracted from the SRTM DEM acquired in Febuary of 2000. A correction is also applied to account for different resolutions between the two DEMs, as the SRTM has a courser resolution at 90 meters (Gardelle et al., 2012). A visual inspection of preliminary results suggests glacier downwasting is occurring in the region (Figure 8). However, statistical error modelling of the Hexagon DEM must first be performed before any real conclusions can be made. This remains an area of ongoing research.

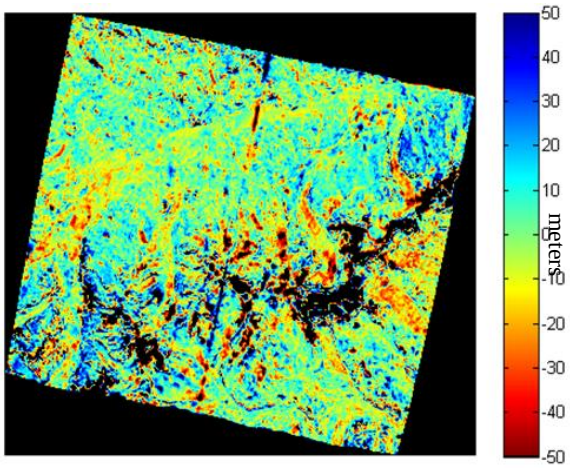

Figure 8. Difference between the 1974 Hexagon DEM and the 2000 SRTM DEM. While some DEM inaccuracies due to cloud cover and poor image contrast over snow-covered regions exist, ice surface lowering over glacier ablation zones is evident, shown as red/orange "tongues" extending both northward and southward (refer to Figure 7 for geographic reference).

CONCLUSION

A new automated workflow for DEM extraction will allow researchers from any discipline to easily and efficiently tap into the vast resource of Hexagon spy imagery. Possible applications include geomorphology and land change studies (glacial retreat, coastal evolution, fault displacement, stream erosion, etc.) among others. The tedious and time consuming process of manual ground control point selection is effectively bypassed, making the Hexagon image database much more appealing, accessible, and applicable. Quantification of Himalayan glacial retreat over several decades is illustrated as a possible research application using the new method.

REFERENCES

Bay, H., Tuytelaars, T., and Van Gool, L., 2006, Surf: Speeded up robust features, in Computer Vision–ECCV 2006: Springer, p. 404-417.

Betts, H.D., and DeRose, R.C., 1999, Digital elevation models as a tool for monitoring and measuring gully erosion: International Journal of Applied Earth Observation and Geoinformation, v. 1, p. 91-101.

Bradski, G., and Kaehler, A., 2008, Learning OpenCV: Computer vision with the OpenCV library: O'Reilly Media, Inc.

Dementhon, D.F., and Davis, L.S., 1995, Model-based object pose in 25 lines of code: International Journal of Computer Vision, v. 15, p. 123-141.

Fusiello, A., and Irsara, L., Quasi-euclidean uncalibrated epipolar rectification, in Pattern Recognition, 2008. ICPR 2008. 19th International Conference on, IEEE, p. 1-4.

Gardelle, J., Berthier, E., and Arnaud, Y., 2012, Impact of resolution and radar penetration on glacier elevation changes computed from DEM differencing: Journal of Glaciology, v. 58, p. 419-422.

Hartley, R., and Zisserman, A., 2003, Multiple view geometry in computer vision: Cambridge university press.

Hirschmuller, H., 2008, Stereo processing by semiglobal matching and mutual information: Pattern Analysis and Machine Intelligence, IEEE Transactions on, v. 30, p. 328-341.

Huggel, C., Schneider, D., Miranda, P.J., Delgado Granados, H., and Kääb, A., 2008, Evaluation of ASTER and SRTM DEM data for lahar modeling: a case study on lahars from Popocatépetl Volcano, Mexico: Journal of Volcanology and Geothermal Research, v. 170, p. 99-110.

James, L.A., Hodgson, M.E., Ghoshal, S., and Latiolais, M.M., 2012, Geomorphic change detection using historic maps and DEM differencing: The temporal dimension of geospatial analysis: Geomorphology, v. 137, p. 181-198.

Kucera, G., 1992, Time in geographic information systems: CRC Press.

Lewis, J., Fast normalized cross-correlation, in Vision interface, p. 120-123.

Perry, A.R., The flightgear flight simulator, in 2004 USENIX Annual Technical Conference, Boston, MA.

Surazakov, A., and Aizen, V., 2010, Positional accuracy evaluation of declassified Hexagon KH-9 mapping camera imagery: Photogrammetric Engineering and Remote Sensing, v. 76, p. 603-608.

Triggs, B., McLauchlan, P.F., Hartley, R.I., and Fitzgibbon, A.W., 2000, Bundle adjustment—a modern synthesis, in Vision algorithms: theory and practice: Springer, p. 298-372.

Wolf, P.R., and Dewitt, B.A., 2000, Elements of Photogrammetry: with applications in GIS: McGraw-Hill New York.